DOCUMENT RESUME

ED 277 249 FL 016 221

AUTHOR Donmail, B. Gillian; And Others
TITLE Evaluation and Assessment. Papers by Members of the Language Awareness Working Party.
PUB DATE 85
NOTE 27p.; In: Donmall, B. G., Ed. Language Awareness. National Congress on Languages in Education Assembly (4th, York, England, July 1984). NCLE Papers and Reports 6; see FL 016 215.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Viewpoints (120) -- Collected Works - General (020)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Course Evaluation; Curriculum Design; Evaluation Criteria; Evaluation Methods; Evaluation Utilization; Foreign Countries; Formative Evaluation; *Introductory Courses; *Language Role; Language Teachers; Linguistics; Secondary Education; Second Language Instruction; Self Evaluation (Groups); *Student Evaluation; Teacher Role; Testing; Test Reliability
IDENTIFIERS Great Britain; *Language Awareness

ABSTRACT

Three papers address issues in course evaluation and assessment of student performance in language awareness programs. The first paper concentrates on the formative evaluation of a course, during its design phase or in progress. It focuses on self-evaluation and two features of course design considered essential to course integrity: explicitness of documentation, and a course management approach that includes a monitoring function. The second paper is concerned with summative course evaluation by an externally administered formal examination. It outlines the questions to be considered in organizing such an evaluation, illustrated by one British secondary school's experience. The third paper examines basic issues and limitations in student testing and the usefulness of student performance assessment as part of the overall course or program evaluation process. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED277249

EVALUATION AND ASSESSMENT

B. G. Donmall
J. M. Sinclair
A. J. Tinkel

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Newson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

FL 016 221

EVALUATION AND ASSESSMENT

B G Donnall, J M Sinclair & A J Tinkel

SECTION ONE - INTRODUCTION

This is a package of short papers which members of the Working Party have assembled as a first response to the need for monitoring of Language Awareness courses. The wide variety of individual initiatives in the subject area makes it very difficult to generalise, and we are anxious not to suggest an orthodoxy so early in the development of Language Awareness. However, it became clear that the Working Party would be dodging its responsibilities if it remained silent on the issue of assessment, so we set about preparing some guidelines. The package is made up as follows:

SECTION ONE:	INTRODUCTION: John Sinclair
SECTION TWO:	PROGRAMME EVALUATION: Gillian Donnall & John Sinclair
SECTION THREE:	FORMAL EXAMINATIONS: Tony Tinkel
SECTION FOUR:	IN-SCHOOL ASSESSMENT OF INTRODUCTORY COURSES: Gillian Donnall

The first point is to separate evaluation and assessment. By evaluation is meant a critical report on a course as a whole - materials, teachers, students, timetable, working conditions, place in curriculum, and anything else that is deemed to be relevant. We believe that Language Awareness courses should strive to be self-evaluating, so that any external evaluation will be straightforward, and, hopefully, positive. That is the professional priority. Language Awareness courses have to justify themselves, and because of their general character they will not be likely to justify themselves on student performance in terminal assessment. Wherever this is possible, of course, we advocate terminal assessment, and Section Three gives useful guidelines for this. But most Language Awareness courses share features which make assessment less than straightforward. They are often of short duration and intended to serve as a basis for further language learning; although dedicated to improvement in student performance, the improvement is of a kind which is very difficult to measure. They are designed to awaken interest, to contextualise language learning, to examine attitudes, to stimulate motivation, and to heighten perception of, responses to and manipulation of language in a wide range of operational situations. Conventional assessment has little to offer in the measurement of factors such as these.

We understand by assessment the measured results of student performance in carefully constructed test conditions. The concern

is with how the student has coped with a given course of instruction. Despite our worries, expressed above, concerning the relevance of this information to Language Awareness courses, we are conscious that the conventions of the educational system pressure teachers to vindicate their courses by some such assessment. Section Four offers suggestions and comments related to this.

The main problem is that retrieval of content may take precedence over the more subtle objectives that motivate the course. A course can be trivialised if its initial objectives are lost sight of, and short Language Awareness courses are particularly vulnerable. A few scraps of information may seem paltry as the only retrievable evidence of what could be major reorientation.

Where it is necessary in a course that some straightforward assessment will be conducted, we urge that this is designed and carried out as a part of the more general self-evaluation of the course as a whole. Then it will be clear to all what part the assessment plays in the overall programme. Section Two offers some principles for setting up a programme evaluation.

John Sinclair

SECTION TWO - PROGRAMME EVALUATION

Introduction

In this paper we shall concentrate on the evaluation of a course while it is in progress - whether it is in the initial stages, under development, or fully up and running. We shall draw mainly from experience with language skills courses; often ESL (English as Second Language) and LA (Language Awareness) but also EFL (English as Foreign Language), EMT (English Mother Tongue) and FLT (Foreign Languages in EMT context).

A thorough evaluation of a course would be ongoing and would slice through and examine the total process at regular intervals. The 'ongoing' cannot be fully separated from the 'before' and the 'after', so we should like to place our remarks first of all with reference to those other points of evaluation, pre- and post-course evaluation.

In general, we think that retrospective evaluation of a course should be designed separately from the course itself. The external evaluator needs to be free from even the most basic assumptions of a course, and the assessment of student performance should also be done in isolation from the course, both to avoid self-fulfilling results, and to avoid the backwash effects of assessment. This separateness will protect the integrity of the course but at the expense of increasing its vulnerability.

So it is to the vulnerability of a course that we turn, and that feature leads back to the very beginning. A course whose design recognises the need for continuous evaluation will be much less vulnerable than one which does not. We find it useless to consider ongoing evaluation without reference to the original design of the course. This kind of evaluation does not need to be external, nor do we need the slightly absurd notion of internal evaluators who are distinct from participants. It becomes a part of the progress of the work, in which everyone participates.

We shall concentrate on two important design features of a course, and show how they lead to reliable self-evaluation. These are:

A. Explicitness of documentation
B. Two - tier management structure, one tier playing a monitor role.

We all get involved in courses, sometimes helping to design or implement them, sometimes to revise and develop, and sometimes just to serve as a team member when a course is well-established. If these two features are present in the design, and properly articulated in the course itself, then the course is fairly safe,

and will stand up to external evaluation at any time. If they are not present, the course is vulnerable, and the people working on it may be subject to unfair criticism.

A. EXPLICITNESS OF DOCUMENTATION

In this section, we must again distinguish between what happens before and after, this time taking a course stage by stage.

1. Forward-looking documentation

In general, the implementation of a course from design to reality requires that we take a series of steps in realisation. The characteristic progress is from abstract to concrete. Each step taken produces a sub-plan which can be checked against requirements in the more abstract plan which stimulated it. So a Language Awareness course might decide on a syllabus item "variety" and then decide what aspect of language variation would be appropriate, and then how available resources could be used to plan a series of lessons, and then what would actually happen in the classroom.

The question here is fidelity of realisation. Each step requires a number of decisions, and the exercise of creative imagination, in how a concept can be brought to life in the classroom. At this point the direction of the course could shift, priorities could subtly change, a misunderstanding could be given licence. Unless the more abstract statement is explicitly made before the step in realisation is taken, the check cannot be made, and the control is lost.

There are additional advantages to an explicit step-by-step design, for example simplifying late stages of implementation like materials-writing. With reference to evaluation, it ensures that the elements of the design are maintained right through the complex process of putting it into practice.

Here is an example. In most language courses of all types, learner motivation is a problem. Despite progressive developments in teaching, it is often difficult to stimulate a class with the prospect of fairly distant goals in order to get them to tackle what is undeniably hard work. There may be few external factors to show them the value of language skills. Hence motivation should be a major design element. If the course is to be successful, each lesson should be interesting or compelling enough in itself. This is a grave design constraint; many teachers may feel that it should figure in every educational course.

If the requirement to provide lesson-by-lesson motivation is to be carried right through, then the specification for each activity must include a statement of the reason why a learner should bother doing it. When the activity is fully articulated, it can be checked

against the motivation statement; when the activity is piloted, it will be discovered if the prediction was correct.

The process of specifying everything step by step in advance is laborious, and can appear unnecessary to experienced teachers in a course team. There is a human tendency to skip the red tape and try something out - a confidence in improvisation which, in its proper place, is essential for teachers who have to survive bad courses. But to avoid bad courses, the discipline of explicit documentation is necessary, and turns out to be highly educational for the participants.

2. Backward-looking documentation

The type and quality of decisions taken in a course should be a matter of record. All sorts of decisions are taken and memory cannot be relied upon. In a course of any size, some individuals may be implementing decisions in ignorance of how and why they were taken.

We shall first of all pick out a few of the ways in which decisions can vary, giving examples from experience.

(a) Decisions can range between principled and arbitrary. Arbitrary decisions are often necessary, and there is nothing wrong with them so long as everyone knows that they are arbitrary and can be varied if good reason arises. For example, in one course, the pilot materials were arranged in four sections of equal size, and the syllabus prescribed one lesson from each section in rotation. Each section became a book which approached the target skills in a different way.

After piloting, this organisation was queried on the balance of the sections (were they equally important?), on the daily switching from one to another (was it unsettling?), and on the distinctiveness of each approach (were they actually separate in practice?). The initial decision to make four sections was a purely operational one, in that there were four senior members of the team, and the most efficient organisation of effort was to divide the job into four parts, each part focusing on a particular priority. The course improved noticeably when three of the four sections were integrated in revision, the placement of lessons next to each other became non-arbitrary, and the balance between sections gave way to a better deployment of the various approaches.

In another project, however, there were and still are four sections because that division was the result of decisions in principle, unrelated to the team structure. Research had shown that the target learners were weak in listening comprehension,

vocabulary richness, and interactive skills, and that remedial attention to syntax was needed for some. These four areas therefore retained their separate identity.

(b) Decisions may be internal or external: external ones are not decisions taken by the course participants, but are delivered from outside. It is very important to record these because they often conflict with internal opinion. For example, it used to be necessary (and sometimes still is) to devise hour-long sessions in the language laboratory, because the booking systems allowed nothing less, and classes which did not use the time fully were placed low in priority.

(c) Decisions may range from theoretical to practical - in the sense of arising from practice. Course designs are all the better for a strong theoretical orientation, but as the implementation proceeds the need for attention to practice becomes stronger, and needs to be accommodated. For example, there once was a course based on formal grammar which introduced words and structures according to the grammar's estimate of their difficulty. The teachers had problems because question forms were regarded as difficult and didn't appear for some time, and information questions were held to be very difficult indeed.

(d) Decisions may range from considered to spontaneous, as contingencies arise. The spontaneous ones must be identified as such so that they can be retrieved for a more searching appraisal at a later stage. Many valuable and innovative features of courses have had their origins in chance or coincidence.

The importance of recording the type and quality of decisions becomes clear when a course is to be revised, adapted or updated, because people develop a fierce loyalty to their creations, and tend to resist the necessary upheavals that keep a course healthy. They may identify with superficial features of the course which can be shown to have little solid support in the design.

Since in an explicit design any feature of the course can be referred to the decisions which originated it, it becomes possible to assess the seriousness of a problem. If changes are proposed, their effect on other parts of the course can be discerned, and disturbance can be kept to the minimum. For example, a good course will try to achieve a balance between theoretical and practical concerns. But once the course is at the piloting stage, and from then on, the feedback is almost entirely from practice, and the influence of theory could disappear altogether. Only if the decision chains are traceable back to source can a theoretical point which is called into question be dealt with at the appropriate level.

In this way a course can remain flexible, and yet consistent with its origins. Flexibility is essential because there are many variables which lie outside the control of the course team. Even fundamental constraints presented as external decisions can turn out to be negotiable, and basic "given" facts can turn out to be wrong. Rigidity is the enemy of good courses, because the problem is changing all the time. One's own course is unlikely to be the only one affecting the total educational pattern. Hence, for example, the expected competence of learners may vary against time, either upward or downward, or change in character. Or the course itself may reveal that an assumption made about the learners is unfounded. A good course will respond to these upsets very economically.

It is worth noting here, with regard to flexibility, that modern technology is making possible the production of materials of high presentation quality, but without the fossilisation of conventional publication. At last, materials can be granted their proper status as by-products of educational courses, and not as the ultimate course goals. With equipment such as word-processors, changes can be made locally in materials to implement revisions, without necessitating extensive overhauls. The advantages to be gained by explicit documentation are thus well worth the labour involved. The course becomes potentially self- evaluating. We turn now to a feature of the management structure which will bring out that potential.

B. TWO-TIER MANAGEMENT STRUCTURE

The suggestion is that each course will set up for itself a monitoring system, which will ensure its relevance from the outset, and give early warning of problems. The first step is to ask who is going to be affected by the course. The full answer will no doubt differ from course to course but typically the list will include

(a) learners
(b) teachers and other staff
(c) institutions, authorities and sponsors
(d) the profession at large.

For each group, we ask the questions - What does the course offer them? What is their method of evaluation? How does the course take account of this?

Within the group, there may be two sets of answers to the questions, one from those who are directly involved in the course and one from those who are only indirectly affected. We must also expect the course to affect groups in ways to which their own evaluation is not sensitive. For example, the question of coverage of a syllabus may not be raised by learner evaluation.

From the concern of interest groups, a report structure and a

committee structure can be devised which will further help the course to be self-evaluating. This monitoring tier should be as fully developed as resources allow, and should be seen to be as important as the teaching. It should not be seen as a necessary chore or a waste of resources or a hindrance to innovation.

Here are some brief examples of the point of view of those affected directly by a course. One avant-garde scheme overseas attempted to teach purely communicatively, but the learners tended to stick to their traditional learning patterns. The learners secretly worked out an underground methodology; for example, there was no rote-learning of vocabulary, so they maintained their own lists.

In some parts of the world teachers do not allocate a lot of time to preparing a lesson based on a text. They expect to be provided with a bank of selected texts, each with a set of conventional exercises following, and they have an efficient, generalised method for getting through the text. They tend to be initially hostile to a course which requires a lot of personal preparation, different for each text. Educational institutions prefer courses to be as standardised as possible, and will examine new proposals to see if they fit into the existing pattern. If there are likely to be discrepancies - for example, a Language Awareness course involving several departments can look anomalous and cause timetable pressure if it takes place in one term only - the course designers can save a lot of trouble by anticipating such a viewpoint, and designing the course so that it fits in.

Information bulletins, etc. are usually welcomed by others in the profession. Not all course organisers are coy about this, but too many are. The slow and patchy spread of information about Language Awareness courses is a case in point.

Let us say a little more about some of these interest groups.

(a) <u>The learners</u>

Learners require a perception of relevance and personal progress to gain confidence in a course. The criterion of relevance can be subdivided - for example, in the area of subject matter. To many staff, the choice of subject matter is regarded as fairly cosmetic, especially when they are working from general principles. To the learner it may be critical, catching a fashion or an interest at just the right time to motivate.

Learners will also be very conscious of the level and tempo of their courses, and time spent in studying their needs will bring dividends. The style of teaching, too, should be carefully matched to expectations. This does not mean that there can be no innovation, but rather that the monitor system constantly checks

that the features of the course are correctly interpreted to learners.

We are led once again, and not for the last time, to a plea for explicitness. Whoever the learners are, we feel an imperative to make sure that they understand the aims and methods of the courses they are taking. Otherwise the opportunity for genuine co-operation in the teaching/learning process may be lost.

This is a unilateral responsibility on the staff side. Many learner groups respond with indifference to initial statements of aims, etc. and retain little understanding of the rationale of the course. There may be many reasons for this, ranging from inability to believe that this level of co-operation and involvement is seriously being sought, to a lack of appreciation of the opportunities offered. When there are difficulties in the way of communication, the effort should be increased.

We have stated this point as forcibly as possible in a short space; obviously it is a design factor like any other, and has to be fitted into the course as a whole. No matter how strongly we feel certain principles, they will be ineffective if they are not integrated into a design which is acceptable to all parties.

Learners must have confidence in their teachers, because through the teachers they will gain confidence in the course as a whole. This is a serious issue in Language Awareness courses, because teachers may be tempted beyond their competence both in their understanding of the principles of language structure, and their control over data, say, from exotic languages.

The learners' perception of progress is another essential feature of a course. Particularly with innovative teaching, we can observe a situation where the students are actually learning well, but do not realise it. This can blunt motivation and lead to confusion about objectives. The course should provide throughout for frequent and regular feedback of an encouraging kind, showing the learners that they are gaining command over language skills which they did not have before, and understanding more about language. This is a particular problem in the area of Language Awareness, which resists simple assessment techniques.

A course is institution-based, but frequently is required to select only some of the learners. A particular group or subject may be picked out, or learners with a particular profile on early assessment. Pilot stages of a course may have to keep the numbers down. It is important to consider the course in the context of the total membership of the institution, and to deal with the interest of both those within and those outside the course. Quite often an excluded group can attack a course because of its very effectiveness, on the reasonable ground that it is only available to

some; equally an included group can perceive a course as adding substantially to their burdens, and attempt to opt out.

(b) The teachers

Very similar considerations apply to teachers on a course, and their colleagues outside. They must have an explicit understanding of the course, and confidence in the management team - confidence in the design, the recruitment of colleagues, and the infrastructure. They must have a clear perception of role. Throughout a course it is desirable that all the teachers are participants in the development of materials, methods and organisation, and the course itself is seen as a valuable training exercise.

For this to be effectively achieved, an important load is again carried by documentation. Initial briefing of new teachers, and orientation in the early stages, will be carefully planned, and at the centre of it will be an up-to-date presentation of the course, at an appropriate professional level. This is a major requirement when the turnover of teachers is anticipated, but is good practice in any course to minimise confusion and to support the day-to-day traffic in information. In a modest course it may be no more ambitious than a file or two, and as well as the big courses we must not forget the individual initiative of a teacher, where design and monitoring matters are no less important, and where discussion with others on planning may enlist their sympathetic support, even though they are not involved in the actual teaching.

Guidance on an organised basis should be continuously available from experienced colleagues. It may seem on the verge of insult to make a point of this, but in large courses, in an ongoing stage, experience is that it is rather rare. New teachers are frequently ushered into the classroom with no support beyond a set of interim materials. Once the first flush of enthusiasm and creativity is over, there is less chance of attention and resources being given to a course, and unless provision is made from the start, then decline sets in. New teachers fail to appreciate the good qualities of the work, and react against it; lack of maintenance leads to ad hoc patching or simple neglect, and in a short time the course becomes an embarrassment. Sadly, we have to record that institutions do not show a strong desire to protect their investments, and each course must set out to be self-perpetuating. This can best be achieved by creating a full range of feedback mechanisms, particularly involving the teachers.

We have mentioned the importance of any course as a training exercise, but we want to make a further plea for opportunities for teachers on an innovative course to have formal professional training. It is a difficult organisational point, particularly in the early years, and teachers can become shackled by their own value to a course. Not only does this increase the isolation of a course

and threaten its up-to-dateness, but it can cause loyal teachers to feel that colleagues who have gone for further training have improved their relative positions in the institution. Teachers must be guaranteed that their participation in a course, with the extra effort that is so often asked of them, will enhance their professional standing. We believe that this point carries implications for those who provide advanced professional qualifications, but the issue cannot be pursued here. Let us instead give a couple of examples to support the general argument that the course should make specific provision for the teachers, and listen to them throughout.

Many courses in recent years have concentrated on the production of materials for the learner, and the word "teacher-proof" has frequently been used. New methodologies, especially with technological innovations, have appeared to reduce the role and importance of the teacher. Coupled with inadequate orientation and guidance, teachers have often become alienated from a course unable to adapt their skills and experience.

A new development is just appearing, where materials are constructed for the teacher, and not immediately for the learner. If this becomes fully accepted, the teacher will re-assume a central role, incapable of being bypassed, and the state of his/her professional awareness and skill will dominate the planning and maintenance of the course. We can only hope that this development will be supported by all parties.

To take a more specific case, it is not uncommon for groups of teachers to be recruited into a course which was not designed to exploit their skills. A clear and recurrent instance of this is the recruitment of native-speaking teachers to a course whose materials are designed for non-native-speaking teachers. We shall not elaborate on the serious problems that arise, but only point out that however much the native speakers are seen as having special attributes, their special needs are just as vital, and should be anticipated, analysed, provided for and supported strongly.

We mentioned the perception of progress when talking about the learners; it is equally important that the teachers have a clear understanding of the way their learners are progressing, and it is a factor which can affect their confidence and morale. It may be the case, although evidence is lacking, that individual progress in the interactive, communicative kind of teaching is as difficult for the teacher to perceive as it is for the learner. The fact that learners are not all learning the same things at the same time, that we easily forget a previous state of incompetence, that solo performance is hardly relevant - these all obscure the subtly growing command of a language, and make it almost impossible to apply simple measures so that the teachers can check their intuitions.

(c) Institutions and authorities

The course is part of the work of an institution, and must relate to it both internally and externally. There are so many variables here, and so many stories to tell, that we shall confine ourselves to broad generalities and a few examples.

First of all, external liaison between the course and those with authority over it must rest on a clear understanding of objectives. Most courses in our experience, are misunderstood by significant segments of the authority structure. Yet understanding is crucial when the course is under way and external priorities change, as they always do. The negotiations that take place to accommodate such changes depend on goodwill, good communications, but above all on detailed, explicit, agreed and understood objectives.

For example, it is almost inevitable that the authority's perception of the role of a course will change as experience grows. Some of the resources allocated may seem attractive for purposes that were not envisaged at the outset; or the course may be under pressure to accept learner groups whose needs are subtly different from those for whom the course was devised. Motives ranging from over-enthusiasm to cynicism may lead to these diversions. The course organisers require an explicit starting point for negotiations, or they may be accused of the most irritating failure of all - failure to do what they did not set out to do.

It is also a fact of life that assumptions made at early stages in a course need to be revised. Some of them are just wrong and since one of the early tasks of a course is to examine all assumptions, their inaccuracy emerges perhaps even in an embarrassing way. Some initial positions are modified by the early experience of the course itself, and some are changed by external factors. Examples: assumptions made about the competence of learners may be proved wrong by the course assessments; assumptions made about the likely rate of progress may be optimistic or pessimistic; assumptions made about learner numbers may be changed by educational policy. Assumptions made about the learners' acquiescence may be challenged by the learners, or may be sabotaged by political changes.

Institutions tend to value stability over experiment, and tolerate innovation only reluctantly and in the short term. Courses are expected to cause a brief flurry and then to settle down, consolidate and work to routine. This is difficult to align with the volatile state of language teaching methodology; changing patterns in the schools invite frequent adaption by the teachers. Such broadly antipathetic positions can become polarised and entrenched, and cause crippling difficulties. One of the recurrent problems is the unwillingness of institutions to allow a continuing element for Research and Development beyond the first stages of a

course. The balancing problem is the opposition of the more creative members of a course team to any sort of compromise or plateau stage in, say, material production.

These considerations point to a course organisation which tries to be explicit about the interface with the institution, authority or sponsor involved, and which tries to allow evaluations in both directions.

Let us turn now to the internal responsibilities of the course with respect to institutions, etc. Factual documentation again takes pride of place; the diary of work, the functions of sections and individuals, the newsy items are all valuable components of what should be very active communication lines. The process of internal documentation and reporting, which we dealt with earlier, can and should lead to the right sort of external documentation.

But it is very rarely found. Usually the two processes are separate, and the external documentation is full of meaningless data, supposedly impressive tabulation, with all the burning issues damped down. Very few internal documents could simply be transmitted outwards, and the course organisers can become information losers, suppressors, distorters and manipulators. This is one of the most intractable problems in the profession, and it seems to arise from faulty perceptions of role, and often also from inadequate resources for proper documentation. Real or anticipated reactions from authority can sometimes be held responsible, but we would also include bad course design, and often a narrowness of outlook of participants. We don't want to appear idealist here, naively believing that truth will always triumph, or that information is neutral; but rather that since these difficulties can confidently be anticipated, and can be seriously debilitating, they should be reckoned as shaping factors in the course from the outset.

(d) The profession at large

It may not seem at first sight that there should be any pressure on the designers or organisers of a course to provide for continuous contact with the profession at large. The individual teachers, of course, will be interested in self-advancement and the more energetic of them will be anxious to attend conferences, write papers, and help in the publication of course materials. Many courses are sympathetic to requests of this nature. Some go farther, and plan stages of presentation to the outside world, using the conference circuit and the journals as vehicles for attracting professional attention.

We believe that one of the biggest dangers facing a language-teaching course anywhere in the world is isolation. Once begun, it

has a tendency to be inward-looking, as the problems strike, as the complexity increases, as the brainchild becomes the obsession of the doting parents. Perhaps because of an imperfect understanding of the course, participants can become very defensive, resisting criticism and losing any analytic detachment. The unique or innovative features of the course are thought of as a professional lifeline.

The enthusiasm and productiveness which this attitude engenders is invaluable, and few if any courses would get far without devotion well beyond the call of duty. Our point is that if contact with colleagues in a wider context is built into a course design it should not inhibit motivation or inspiration but will remove the need for secrecy and defensiveness. Any good course should be spurred on by regular positive feedback and constructive criticism. There should be a dialogue from the start with other groups and individuals, seeking views and suggestions. It has to be efficiently organised because of the turnround time of comments, but it allows the course team to identify at any time the professional importance of their work, encourages them and others to draw out the likely lessons for other courses, and reduces the depressing delay that we usually have to suffer before good work becomes available. Properly planned, it is a major feature of ongoing course evaluation.

Some courses devise a halfway house for professional contact in the shape of an Advisory Group or the like. That is better than nothing, but not very satisfactory. For one thing, the members of such a group tend to become as inward-looking as the team itself, and for another, they rapidly become sophisticates and therefore unrepresentative. One of the very first courses of a communicative nature, back in the sixties, had really magnificent support from its representative advisers, but was still too innovative when it appeared on the market. The hardbitten heads and their staff were just not ready for it, and the advisers had apparently acted as a buffer from reality rather than a route through to it.

In conclusion, we should like to summarise the central theme. Evaluation is always a sensitive issue, and it can be faced without fear only if courses are properly designed and fully documented. The documentation, while laborious, has a wide range of uses beyond course control. A feature of the design of any course should be a monitoring system so that all groups affected by the course will be conducting an efficient two-way dialogue with the course team. When a course is safeguarded in this way, then external evaluation is no problem, and if possible external evaluations should be based on different principles from those which have guided the course.

Gillian Donnall and John Sinclair

SECTION THREE - FORMAL EXAMINATIONS

1. This section of the paper on Evaluation and Assessment is concerned with terminal assessment of a Language Awareness course by means of an externally administered formal examination. The questions that need to be considered when setting up such an examination are defined through describing the experience gained at The Oratory School in setting up the GCE A/O Level Principles of Language examination. This experience is, of course, based only on groups of able A level students who had elected to follow the course. There are, however, certain fundamental questions that should be faced by anyone setting up any formal assessment at any level.

2. Why is formal external assessment desired? In the case of The Oratory the answer was fourfold.

i) to show that is was possible and feasible to set up an examination in Language Awareness through a mainstream examination system. (see Section 5)

ii) to enhance the validity of the course in the minds of the students. The success of the Graded Objectives and Tests movement has shown that motivation is increased if the course can be seen by the students to culminate in demonstrable and tangible achievement.

iii) to provide a stimulus and a guide for the better internal functioning of the course, as, for example, in the case of the extended essay element of the examination. The preparation of this 2500 word essay on a subject of their choice made the students think more deeply in an aspect of the course than would otherwise have been the case.

iv) to help define and confirm the nature of the course in the minds of teacher, students and onlookers. The content of any examination paper and the way in which the questions are framed have a powerful influence on people's view of the syllabus and how it should be taught.

3. What is the nature of the course being assessed?

i) Before starting to plan the form that an external examination is to take, it must be decided what the course is meant to achieve in those who follow it. In the case of The Oratory the course aims were:-

a) to make the students more sensitive to their own and people's use of language, and in the process
b) to provide a pre-theoretical introduction to the basic principles of linguistic study.

ii) It must also be decided what means are to be used to realise the aims of the course. The syllabus content of the course at The Oratory is detailed in the next section of this publication, Reports and Initiatives in Core Schools. Briefly it covers:-

a) definition of human language and communication
b) analysis of English
c) examination of change and variety in usage.

iii) It must be considered carefully whether the objectives of the course will lead to the aims being realised. Will the syllabus have the intended effect on the students? In the case of The Oratory this question becomes:- Would coverage of the syllabus in 3.ii develop the students' awareness of what language is and how it functions? Would that increased awareness of language produce increased sensitivity to how it is used? Unless the probable answer to such questions is positive, the course is not linked to its stated aims.

iv) The way in which the course content is presented to the students is also important. Does the teaching of the syllabus suit the content of the syllabus? Do the manner of exposition and the format of the materials lead the students effectively towards achieving the aims of the course? The methodology of the course at The Oratory is based on the fact of the students' native speaker knowledge of English. It is felt that the best way of making the students more aware of language is by guiding them to explore what they intuitively know and use already. The students are guided towards an awareness of linguistic concepts primarily through the analysis of examples of that concept. If the process is successful, then it should lead on to their producing further examples of their own. (Further discussion of this methodology can be found in the paper Methodology Related to Language Awareness Work.)

4. What form should the examination take?

i) If the form of the examination does not correspond to the nature of the course, then the nature of the course will change to fit the form of the examination. Conversely, the aims, objectives and methodology of the course are reinforced, if the form of the examination coincides with them. If an examination terminates the course, the teaching will inevitably be directed towards the form and content of the examination.

ii) The content of a course may be primarily concerned with increasing students' knowledge, or alternatively with increasing their skill and performance. The course at The Oratory is of the latter type, since it aims at increased sensitivity to language use through developed awareness of its nature and function. The questions in the examination had, therefore, to reflect the fact

that it was performance in being aware of language that was to be tested, not knowledge of how to talk about language.

iii) More specifically the questions in the A/O Principles of Language examination had to fulfil a twofold function:-

a) to ascertain whether there was an improvement in the awareness of the student about language. Such improvement was deemed to have been shown by the satisfactory completion of the linguistic tasks in the questions;

b) to state the linguistic tasks in such a way that individual insight on language could be shown, either in the form of an unexpected slant on the task in hand or of seeing wider implications.

iv) The format of the questions had to replicate the methodology of the course at The Oratory. The emphasis in the course presentation was on guided exploration through examples, on the students' approaching analysis and definition through handling examples of language for themselves. The examination questions were framed to fit this pattern, thereby reinforcing it. (For examples of the questions see Report of Work at the Oratory School, Section 7.)

v) If the Principles of Language examination was restricted to a question paper composed of language examples presented to the candidate for handling (as outlined in Section 4. iv), then they would have little opportunity for continuous writing in the context of the examination. This objection was rectified by introducing an extended essay into the examination. The students were required to write about 2,500 words on a topic of their own choosing related to the syllabus. This essay would be researched and written over a period before the question paper was sat. It gave the students the stimulus to investigate more closely an aspect of the course that interested them, thereby reinforcing the objectives and aims of the course by another means. (Topics chosen by candidates for their extended essays are listed in Report of Work at the Oratory School, Section 8).

vi) When the form of the terminal examination has been decided, the results of the first sittings must be monitored to see whether the examination is doing what it is intended to do. For example:-

a) Do the results coincide sufficiently well with the pre-examination predictions of the teachers? Is there a good spread of results? Have the better candidates been given the opportunity to shine?

b) Do questions produce the kind of answers expected? Do questions have equal value or is it easier to gain credit by answering some questions in preference to others? If it is not already the case, should some questions be made obligatory and should the paper be

sectionalised? Is the time limit too generous, too short or about right? Is the lay-out of the paper a help or a hindrance to the candidates? Are the overall examination instructions clear? Is there any question rubric which is misleading? Could the candidates have been expected to know the terminology used? If not, was such terminology explained?

c) If any component of the examination has been completed before the written examination, such as the extended essay referred to in Section 4.v, has the opportunity to work in greater depth been taken up by the candidates? If not, is this the fault of the terms of reference of this part of the examination?

d) If the candidates have been taught by one teacher or a small group of teachers in the same school, what problems would be involved in opening up the examination to a wider public? How far do the teaching on which the examination is based and hence the preparation for that examination depend on particular teachers in particular circumstances? Can the syllabus, method and materials be taken up by other teachers in other schools and used in their individual contexts, allowing for the fact that they can still make personal contact with the original source in case of difficulty? Can they be taken up more widely when the only feasible contact between participating schools is that of a formal statement of syllabus content and formal instructions on examination procedure?

e) What adjustments would have to be made to the arrangements for setting, marking, moderating and administrating the examination, if it was offered to a wider public?

5. i) It is always possible to approach the ruling body of an examination system with proposals for a formal examination under its aegis. In the case of the Oratory's Principles of Language examination the [illegible] with the Oxford and Cambridge Schools Examination Board [illegible] statement of aims and objectives, an outline of the course syllabus, an indication of the intended age and ability level of the candidates and an outline of the proposed examination. When the Board decided to investigate the proposal, it asked for further information in the form of a specimen examination paper, detailed guidelines on how it would be marked and a marking scheme. Detailed consultations produced a format that the Board felt was sufficiently consistent with the aims and objectives of the original submission, as well as with their own examining standards, for the proposal to be given a three-year trial period restricted to candidates from the place of submission. In this way the proposal could be tested in practice, free from the problems that would occur in sharing an experimental course and examination between schools.

ii) [illegible] an examining body for a formal

examination, attention must be given to the financing and marking of the proposed examination. How will the costs of the examination be covered? Who will set, check and produce the question papers? Who will mark and moderate its various components? How far will the teachers of the course be involved in the setting and marking of the question papers? An examination body will offer guidance and opinion on these questions which will vary according to the nature of the board, depending in particular on how teacher-based it is. Such mundane questions are crucial, since their answers will decide the smoothness of the routine administration of the new examination and thereby facilitate its being offered to a wider spread of candidates, if and when the initial pilot testing is successfully completed.

iii) A formal examination must respond to expectations of what a formal examination is like. If the reality falls short of such expectations, the formal assessment will not be seen as a serious examination and will not be treated as such by students on the course, other teachers or members of the public. If the underlying aims and objectives are thoroughly thought out, if the examination itself is properly planned and prepared in the context of those aims and objectives, if the presentation is that of a serious public examination, then it will evoke a serious response.

A.J. Tinkel

SECTION FOUR: IN-SCHOOL ASSESSMENT OF INTRODUCTORY COURSES

(It is appreciated that some of the following will not be new to many teachers, that some of the points made are applicable to all testing situations. However, they are offered as a composite source for reference with the particular circumstances of Language Awareness initiatives in mind).

A number of the points made in Section Three of this paper with regard to the setting up of a formal assessment scheme with an outside examining body are as important to any internal assessment scheme. Few teachers can be unaware of the problems of setting tests which successfully do what is required of them; testing is no easy matter even for experts. Few teachers are given thorough training in the setting of tests/examinations and yet the results of internal assessment procedures are often crucial determinants of the shape of a pupil's further learning. For example, as a result of an internal examination a pupil may be excluded from any further learning of the subject or allocated to a specific learning level. Alternatively, poor performance in a test may have a strongly negatively reinforcing effect on his/her learning of that subject if inadequacy rather than achievement has been underlined.

It should be said again, as was first pointed out in Section I, that we do not favour formal assessment procedures for Language Awareness work, that on-going evaluation should be a much more effective monitor of a course, its teachers and its learners. Since most Language Awareness courses in existence at present are of relatively short duration, not destined for qualification as O- or A-level subjects, since they are designed to further competence in learning in all curricular areas, since emphasis is on skill and performance, not factual knowledge, and that is per se more difficult to assess with a measure of objectivity, assessment as a means of determining a pupil's progress (and thereby the value of the course as a whole) is not especially to be recommended.

Foreign language teachers are only too well aware of examination techniques which in order to attempt a measure of objectivity and allow for standardization incorporate conventions of examining far removed from the language use in most schools' teaching syllabuses. Furthermore, they are well aware of the "backwash effect" of examining techniques on teaching methods.

Where assessment is still required, however, we should wish to emphasise that it should form only a part of a whole evaluation process within which all work produced plays an important part in the profile of a pupil's learning and development, thus ensuring that conclusions are not drawn from an assessment undertaking which it cannot logically be expected to provide.

Why ?

For whom or for what purpose is a formal assessment process to be carried out?

If it is sought by the head, for other staff, for parents, as a means of justification of the course in terms of improved pupil conscious awareness and performance, so be it, but its place in the overall evaluation process should be underlined. It should not stand alone as the sole determinant of the value of the course defined in these terms.

If the response is "for the pupils", the teacher might question that necessity. It should be a built-in on-going feature of a Language Awareness course that learners perceive and give evidence of their own progress.

If the answer is : to use as a means of discriminating between pupils in order to determine placement, for example, into ability sets, this would run counter to the philosophy underlying all Language Awareness work. The intention with such a test, is to arrive at a spread of marks, whereby a substantial number of pupils are labelled as "less competent". However, in Language Awareness work the basis is pupils' already high performance skills in language; its purpose is to make conscious and develop those skills to higher levels. We should intend any assessment of pupils in Language Awareness to record levels of achievement not inadequacy.

When ?

Assessment is usually sought at the end of a teaching programme and we should not wish to recommend any greater frequency. Diagnosing and putting right pupils' problems and ensuring continued progress will form part of the teaching programme of a Language Awareness course which is subject to continuous monitoring.

What ?

a) The essential setting-off point is : the aims and objectives of the course, the objectives of each teaching module within the course and of each small section within the module, firstly in order for goal-orientated teaching to take place and secondly for proper assessment of it to be tackled.

e.g.

Aim:
Appreciation of the variety and range of expression present in the United Kingdom and the circumstances and appropriacy of their use.

Objective:
Knowing to which part of the country different expressions belong through familiarity with their use in speech.

Assessment:
Listen to recordings of speech and indicate on the map provided in which part of the country they are widely used.

b) All Language Awareness courses have an affective as well as a cognitive aspect (see The Report, section: The Term Language Awareness). This aspect in itself defies assessment but it would be wrong to overlook it when seeking a view of the value of a course as a whole. This is an important part of the evaluation.

c) As the core school descriptions show, most, if not all, Language Awareness courses have some input of knowledge content. This can be a relatively simple kind of area to assess if extension of knowlege in certain defined fields is a complete course objective in itself. Pupils may, for example, be required to reproduce that knowledge on the stimulus of a request for the same.

e.g.

In which countries is German/Greek/Spanish spoken?

However, it is important to remember that knowledge content in Language Awareness courses is usually subservient either to the affective sphere or as a preliminary to and ultimately part of the conscious awareness, skill/performance sector. In these circumstances the knowledge aspect should not be assessed separately but as part of the whole activity which reflects the course objective.
e.g.

Knowledge of intonation patterns allied to awareness of how they affect meaning, may allow a pupil, for example, a. to say a sentence with at least 4 different stress patterns to indicate 4 different meanings ("we'll paint the garage red and white as well") and b. to state what the different meanings conveyed are.

d) In Language Awareness work two important concerns are appropriacy and effectiveness of language use. A pupil might be asked to indicate the most appropriate response in given circumstances.

e.g.

Situation : pupil late for athletics training because he/she had been chatting to his/her friend.

Possible comments to the teacher: (a) I'm very sorry, (b) I'm

stricken with mortification, (c) I do apologise profusely, and to give reasons for his/her choice.

Or again, a pupil might be asked to express him/herself as effectively as possible, e.g. persuade a customer to buy fruit from your stall in a market.

Objectives related to an appreciation of styles and registers of language may invite an assessment item of the following kind:

Situation: man returns home after work to find that his house has been burgled. Many valuable items have been taken and also his guard dog!

Task : quote briefly his description of what has happened to the following people : (a) his wife when she comes home from work, (2) the policeman who comes to investigate (3) the postman who delivers the mail early next morning (4) his friends in the pub. It will be appreciated that assessment of these, especially the latter, types is not easy but the assessment must respond to the objectives, the objectives must not be lost in the interests of standardization or objectivity.

How ?

a) Syllabus:

The teaching syllabus and methodology allied to aims and objectives must be detailed as an essential factor in the design and implementation of a Language Awareness course. Any assessment procedure must relate, as best as it can, to these. It is not acceptable for limitations or conventions of assessment to condition either course content or teaching methodology.

b) Validity:

Not only must an assessment process reflect the content and teaching methodology of the course but every item must actually achieve the intention assigned to it in order for it to be valid. For example, if pupils have been learning to discriminate between regional speech patterns based on listening to examples on tape, it is not valid to assess their competence by giving them examples on paper!

c) Reliability:

The larger the number of items, the more likely the test is to be reliable, i.e. the same statistical pattern would again be produced. In any case it would be unsatisfactory to use the single instance to make a statement about the generality of a pupil's progress.

d) Balance:

It is important that the balance of skills to be assessed reflects the balance determined for them in the course design. It is only too easy, for example, to include a number of items which are easy to assess, such as those related to extension of knowledge, and then to weight them in accordance with that, although the course design and purpose gives only small value weighting to this area per se.

e) Instructions:

Not only should these be absolutely clear but should be familiar to the pupils from classroom teaching use.

f) Time allocation:

The intention is to assist pupils to give the best account of themselves in circumstances most conducive to that. We should not wish to advocate that speed of writing, for example, should also be assessed. For this reason adequate time should be allowed for the tasks set, or if the constraint of time has to be the first priority, the tasks must be limited to suit those constraints.

Marking

a) All Language Awareness work is based on achievement. Pupils already have high performance skills and the intention is to advance these together with their conscious perception of language in use. Tests therefore should show what the pupils have achieved, not what they have not. The test will be set in such a way as to reflect this and thereby give stimulus to pupils' further learning. The philosophy underlying the Graded Proficiency test scheme is important to the Language Awareness field.

b) Appropriate marking must reflect the defined objectives of both teaching and testing, e.g. if correct spelling is not a defined objective, it should not be marked. The determining of parameters is always difficult but examiners must discipline themselves to mark only what it was intended to assess or otherwise run the risk of invalidation. (Consideration may be given to additional marks for especially good work in defined areas).

Evaluation

In the case of evident success, few problems arise. In the case of lack of success this may be due as much to, for example, teaching methodology, materials, course design, assessment design as to pupils' incompetence or laziness or the impossibility of achieving positive results in Language Awareness work. A proper evaluating

and self-amending programme should ensure avoidance of most, if not all, of these but it is important to realise that an assessment process itself should also be subject to evaluative scrutiny. The following sample questions may give guidance for this:-

- do the results coincide with predictions and evidence given in class work?
- have all ability levels had the opportunity to show what they can do?
- was it possible to "play the system", i.e. by choice of some questions in preference to others?
- were the kind of responses intended actually supplied?
- was time allocation reasonable?
- was it a positive undertaking?
- did the marking scheme reflect the balance of objectives?
- what was the customers' response? i.e. what were their views of the assessment procedure?

In conclusion, the limitations of assessment processes and the problems with which they are beset underline how important it is that results obtained from them should not be used as the sole determinants either of pupils' achievement or of the value of a Language Awareness course. Section IV is offered to assist teachers to assess as effectively as possible when this is necessary but we hope that it serves also to emphasise the need for an on-going evaluation programme of which any assessment undertaken would simply form a part.

Gillian Donnell